DECLARATION

I, the undersigned, of 15-29, Tsukamoto, 3-chome, Yodogawa-ku, Osaka 532-0026, JAPAN, hereby certify that I am well acquainted with the English and Japanese languages, that I am an experienced translator for patent matter, and that the attached document is a true English translation of

Japanese Patent Application No. 11-329853.

I declare that all statements made herein of my own knowledge are true, that all statements on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Signature:

Natsuko Honjo

Natsuko Honjo

Dated: November 2, 2006

[Name of the Document] Specification

[Title of the Invention] Image processor, a method of providing image processing services and order processing method

[Claims]

[Claim 1] An image processor including a display for presenting an image of an object thereon, characterized by comprising:

an image synthesizer for generating a scale image representing a substantially real size at a position specified on the image presented on the display in accordance with three-dimensional positional information of the object captured in the originally presented image and for combining the scale image with the image of the object,

wherein a synthesized image obtained by combining the scale image with the object image is presented on the display.

[Claim 2] The image processor of Claim 1, characterized by comprising:

an imaging section for capturing the object image containing the three-dimensional positional information; and

a range image generator for obtaining the three-dimensional positional information from the image captured by the imaging section,

wherein the image synthesizer generates the scale image

in accordance with the three-dimensional positional information obtained by the range image generator.

[Claim 3] The image processor of Claim 2, characterized in that

the imaging section comprises light-emitting means for projecting light with a predetermined radiation pattern onto the object and captures the object image containing the three-dimensional positional information by receiving the light that has been projected onto, and then reflected from, the object.

[Claim 4] The image processor of Claim 1, characterized by comprising an imaging section having an automatic or manual focusing controller,

wherein the image synthesizer generates the scale image by using data that represents a distance of the object and is obtained by the automatic or manual focusing controller as the three-dimensional positional information.

[Claim 5] The image processor of Claim 1, characterized in that

the scale image represents a shape of a ruler.

[Claim 6] The image processor of Claim 1, characterized by comprising input means that is so formed as to allow a user to externally input the specified position.

[Claim 7] The image processor of Claim 6, characterized in that

the input means is a touch panel formed on the surface of

the display.

[Claim 8] The image processor of Claim 6, characterized in that

the input means is a penlike pointing device that is so formed as to allow the user to specify arbitrary coordinates on the surface of the display.

[Claim 9] The image processor of Claim 6, characterized in that

the input means is a cursor key, mouse or press button that allows the user to move a cursor presented on the display and to specify coordinates representing the location of the cursor.

[Claim 10] An image processor including a display for presenting an image of an object thereon, characterized by comprising:

an image synthesizer for combining respective images of multiple objects together in accordance with three-dimensional positional information of each of the objects captured in each of the originally presented images so as to meet a desired size relationship,

wherein a synthesized image obtained by combining the multiple images together is presented on the display.

[Claim 11] The image processor of Claim 10, characterized in that

the image synthesizer combines the image of one of the

objects which has been separated from a background image with another background image.

[Claim 12] The image processor of Claim 11, characterized in that

the image synthesizer cuts out from the original image an image portion that is made up of pixels at respective locations associated with distances falling within a predetermined range as the object image separated form the background.

[Claim 13] The image processor of Claim 10, characterized by comprising:

an imaging section for capturing the object images containing the three-dimensional positional information; and

a range image generator for obtaining the three-dimensional positional information from the images captured by the imaging section,

wherein the image synthesizer combines the images together in accordance with the three-dimensional positional information obtained by the range image generator.

[Claim 14] The image processor of Claim 13, characterized in that

the imaging section comprises light-emitting means for projecting light with a predetermined radiation pattern onto the at least one object and captures the object image containing the three-dimensional positional information by

receiving the light that has been projected onto, and then reflected from, the object.

[Claim 15] The image processor of Claim 10, characterized by comprising an imaging section having an automatic or manual focusing controller,

wherein the image synthesizer combines the images together by using data which represents distances of the objects and is obtained by the automatic or manual focusing controller as the three-dimensional positional information.

[Claim 16] The image processor of Claim 10, characterized in that

the image synthesizer is so formed as to scale up, scale down or rotate at least one of the multiple images.

[Claim 17] The image processor of Claim 10, characterized by being so formed as to allow a user to externally define or change relative positions of the images being combined.

[Claim 18] An image processor including a display for presenting an image of an object thereon, characterized by comprising

an image synthesizer for generating an image representing the object in its real size when presented on the display by scaling the image up or down in accordance with three-dimensional positional information of the object captured in the originally presented image,

wherein the image representing the object in its real size is presented on the display.

[Claim 19] A method of providing image processing services, characterized by comprising the steps of:

receiving an image containing three-dimensional positional information of an object and an instruction of a desired image processing mode from a user of the services;

obtaining the three-dimensional positional information of the object captured in the image from the object image received;

performing image processing on the received image in the instructed image processing mode in accordance with the three-dimensional positional information obtained; and

sending out image data obtained as a result of the image processing to the user of the services.

[Claim 20] An order processing method characterized by comprising the steps of:

receiving an instruction which specifies a type of a product from a user;

sending out an image containing three-dimensional positional information of the product of the specified type to the user; and

receiving an order for the product, along with information specifying the size of the product, from the user.

[Detailed Description of the Invention]

[Technical Field to which the Invention Belongs]

The present invention relates to image processing technology. Particularly, the present invention belongs to a technique of processing an image by utilizing information about the three-dimensional position of an object.

[Prior Art]

Figure **8** is a diagram illustrating a basic arrangement for a rangefinder that can capture a range image (or depth image) (Japanese Unexamined Patent Publication No. 11-144097). In Figure **8, 51** denotes a camera, **52a** and **52b** denote multiple light sources, **55** denotes a light source controller for controlling the light-emitting operations of the multiple light sources **52a** and **52b**, and **56** denotes a distance calculator for generating a range image from an image captured by the camera **51**. In response to a vertical synchronizing signal supplied from the camera **51**, the light source controller **55** gets each of the light sources **52a** and **52b** alternately flashed.

Figure **9(a)** is a diagram illustrating exemplary configurations for the light sources **52a** and **52b**. As shown in Figure **9(a)**, the light sources **52a** and **52b** employ, for example, the configurations in which flash light lamps **57** and **58** such as xenon flash lamps are vertically aligned and reflectors **59** and **60** are placed behind these lamps **57** and **58**, respectively, so as to differ in the horizontal orientation from each other.

Figure **9(b)** is a plan view of the configurations shown in Figure **9(a)**. The light sources **52a** and **52b** radiate light beams within the ranges **A** and **B** shown in Figure **9(b)**, respectively. The xenon lamps used in this case have so small light-emitting portions that these lamps can be regarded as point light sources in the plan view. Also, the light sources **52a** and **52b** are spaced apart from each other by about 1 cm. Accordingly, the light may be regarded as being emitted from almost a point.

The operating principle of the rangefinder shown in Figure **8** will be described with reference to Figures **10** through **13**.

Figure **10** is a diagram illustrating exemplary light patterns that have been radiated from the light sources **52a** and **52b** shown in Figure **9**. In Figure **10**, the solid lines **La** and **Lb** represent the brightness on a screen surface where the light beams have been projected on a virtual screen **Y** from the light sources **52a** and **52b**. The degree of the brightness is represented by how high the solid lines **La** and **Lb** are in the direction indicated by the arrow. As can be seen from Figure **10**, the light projected from each of these light sources **52a** and **52b** has a characteristic that it is brightest on the center axis of the projection range and gets gradually darker toward the edges of the range. This characteristic results from the disposition of the semi-cylindrical reflectors **59** and

60 behind the flash light lamps **57** and **58**, respectively. And depending on which directions these reflectors **59** and **60** face, the light beams projected from the light sources **52a** and **52b** may or may not partially overlap with each other.

Figure **11** is a graph illustrating a relationship between the angle ϕ of the projected light as measured in the direction **H** shown in Figure **10** and the light intensity. The direction **H** is defined as a direction in which an arbitrary plane **S**, including the center of the light source and lens, and the virtual screen **Y** intersect with each other. The angle ϕ is an angle formed by the light, which has been projected onto the XZ plane, with the X-axis. As shown in Figure **11**, when seen from the light sources, a range α of the light pattern projected from one of the light sources **52a** and **52b** through the object's space has bright and dark parts on right- and left-hand sides, respectively, while the range α of the light pattern projected from the other light source **52b** or **52a** through the object's space has dark and bright parts on right- and left-hand sides, respectively. It should be noted that the light patterns shown in Figure **11** change in the height direction (i.e., the Y direction). In other words, the patterns are changeable depending on the level at which the plane including the centers of the light sources and lens is located.

Figure **12** is a graph illustrating a relationship between

the angle ϕ of the projected light and the light intensity ratio in the range α shown in Figure **11.** A relationship between the light intensity ratio and the angle ϕ in the range α is one-to-one.

In this embodiment, to measure the distance of an object, two types of light patterns should be alternately projected onto a plane standing vertically at a predetermined distance from the light source and light beams reflected from the plane should be imaged at the camera **51**. A relationship between the light intensity ratio and the angle of the projected light such as that shown in Figure **12** should be obtained in advance for each Y coordinate (corresponding to a Y coordinate on the CCD). And if the light sources **52a** and **52b** should be disposed so that a line segment connecting the center of the lens of the camera **51** to the light sources **52a** and **52b** is parallel to the X-axis of the CCD imaging plane, then the distance of the object can be estimated accurately based on the data representing the relationships between the light intensity ratios and angles of the projected light associated with the respective Y coordinates obtained beforehand.

Now, take a look at the point **P** shown in Figure **8**. The intensity ratio of the light beams projected from the light sources **52a** and **52b** onto the point **P** is obtained from the image captured by the camera **51**. And the angle ϕ corresponding to the point **P** as viewed from the light sources

52a and **52b** can be derived from the obtained intensity ratio of the light beam by reference to the relationship shown in Figure **12** associated with the Y coordinate of the point **P**. Also, the angle θ corresponding to the point **P** as viewed from the camera **51** can be determined based on the coordinates of a pixel associated with the point **P**, focal length, and various camera parameters including coordinates of the optical center of the lens system. Then, the distance of the point **P** is estimated by the principle of the triangulation using the two angles obtained θ and ϕ and a baseline length, i.e., the distance between the position of the light sources **52a** and **52b** and the optical center of the camera **51**.

Suppose the optical center of the camera **51** is defined as the origin, the direction of the optical axis of the camera **51** is defined as the Z-axis, and horizontal and vertical directions are defined as the X- and Y-axes. An angle which the direction of the point **P** when viewed from the light sources **52a** and **52b** makes with the X-axis is defined as ϕ. An angle which the direction of the point **P** when viewed from the camera **51** makes with the X-axis is defined as θ. And if the coordinates of the light sources **52a** and **52b** are (0, -D), i.e., the baseline length is D, then the depth **Z** of the point **P** is given by

$$Z = D\tan\theta\tan\phi / (\tan\theta - \tan\phi)$$

Alternatively, all of the three-dimensional coordinates (X, Y,

Z) may be calculated by the following equations using the angle ω shown in Figure **13**:

$$X=Z/\tan\theta$$

$$Y=Z/\tan\omega$$

Furthermore, an ordinary color image can also be obtained by adding and averaging the images formed by flashing the light sources **52a** and **52b**. Accordingly, an image containing three-dimensional positional information can be captured by using the arrangement shown in Figure **8**.

[Problems that the Invention is to solve]

However, even though it is now technically possible to capture an image with three-dimensional positional information as described above, it does not necessarily lead to attractive products that will be welcomed by the market. Accordingly, in developing products, it is very important to add functions that are highly convenient and entertaining for users to the products based on the above-described technique.

It is an object of the present invention to realize functions that are highly convenient and interesting for the users by utilizing three-dimensional positional information of an object in image processing.

The present invention also provides a method of providing image processing services and a method of processing an order by utilizing three-dimensional positional information of an image.

[Means for Solving the Problems]

To solve the above problems, a solution taken by the invention of Claim 1 is an image processor including a display for presenting an image of an object thereon. The image processor comprises an image synthesizer for generating a scale image representing a substantially real size at a position specified on the image presented on the display in accordance with three-dimensional positional information of the object captured in the originally presented image and for combining the scale image with the image of the object, wherein a synthesized image obtained by combining the scale image with the object image is presented on the display.

According to the invention of the Claim 1, a scale image representing a substantially real size is combined at a specified position with an image originally presented on the display so that a synthesized image is presented on the display. Thus, the user can know a real size of the object imaged quickly on looking at the image presented.

According to the invention of Claim 2, the image processor of claim 1 comprises an imaging section for capturing the object image containing the three-dimensional positional information; and a range image generator for obtaining the three-dimensional positional information from the image captured by the imaging section, wherein the image synthesizer generates the scale image in accordance with the

three-dimensional positional information obtained by the range image generator.

According to the invention of Claim 3, the imaging section in the image processor of claim 2 comprises light-emitting means for projecting light with a predetermined radiation pattern onto the object and captures the object image containing the three-dimensional positional information by receiving the light that has been projected onto, and then reflected from, the object.

According to the invention of Claim 4, the image processor of claim 1 comprises an imaging section having an automatic or manual focusing controller, wherein the image synthesizer generates the scale image by using data which represents a distance of the object and is obtained by the automatic or manual focusing controller as the three-dimensional positional information.

According to the invention of Claim 5, the scale image generated by the image processor of Claim 1 represents a shape of a ruler.

According to the invention of Claim 6, the image processor of Claim 1 comprises input means that is so formed as to allow a user to externally input the specified position.

According to the invention of Claim 7, the input means in the image processor of Claim 6 is a touch panel formed on

the surface of the display.

According to the invention of Claim 8, the input means in the image processor of Claim 6 is a penlike pointing device that is so formed as to allow the user to specify arbitrary coordinates on the surface of the display.

According to the invention of Claim 9, the input means in the image processor of Claim 6 is a cursor key, mouse or press button that allows the user to move a cursor presented on the display and to specify coordinates representing the location of the cursor.

A solution taken by the invention of Claim 10 is an image processor including a display for presenting an image of an object thereon. The image processor comprises an image synthesizer for combining respective images of multiple objects together in accordance with three-dimensional positional information of each of the objects captured in each of the originally presented images so as to meet a desired size relationship, wherein a synthesized image obtained by combining the multiple images together is presented on the display.

According to the invention of Claim 10, multiple images are combined together so as to meet a desired size relationship, e.g., so that the images are substantially in proportion to their actual sizes, and then a resultant synthesized image is presented on the display. Thus, for

example, the user can combine an image of the object with another background image with their scales matched to quickly watch a virtual image of the object on the different background image.

According to the invention of Claim 11, the image synthesizer in the image processor of Claim 10 combines the image of one of the objects which has been separated from a background image with another background image.

According to the invention of Claim 12, the image synthesizer in the image processor of Claim 11 cuts out from the original image an image portion which is made up of pixels at respective locations associated with distances falling within a predetermined range as the object image separated from the background.

According to the invention of Claim 13, the image processor of Claim 10 comprises an imaging section for capturing the object images containing the three-dimensional positional information; and a range image generator for obtaining the three-dimensional positional information from the images captured by the imaging section, wherein the image synthesizer combines the images together in accordance with the three-dimensional positional information obtained by the range image generator.

According to the invention of Claim 14, the imaging section in the image processor of Claim 13 comprises light-

emitting means for projecting light with a predetermined radiation pattern onto the at least one object and captures the object image containing the three-dimensional positional information by receiving the light that has been projected onto, and then reflected from, the object.

According to the invention of Claim 15, the image processor of Claim 10 comprises an imaging section having an automatic or manual focusing controller, wherein the image synthesizer combines the images together by using data which represents distances of the objects and is obtained by the automatic or manual focusing controller as the three-dimensional positional information.

According to the invention of Claim 16, the image synthesizer in the image processor of Claim 10 is so formed as to scale up, scale down or rotate at least one of the multiple images.

According to the invention of Claim 17, the image processor of Claim 10 is so formed as to allow a user to externally define or change relative positions of the images being combined.

A solution taken by the invention of Claim 18 is an image processor including a display for presenting an image of an object thereon. The image processor comprises an image synthesizer for generating an image representing the object in its real size when presented on the display by scaling the

image up or down in accordance with three-dimensional positional information of the object captured in the originally presented image, wherein the image representing the object in its real size is presented on the display.

A solution taken by the invention of Claim 19 is a method of providing image processing services, comprising the steps of: receiving an image containing three-dimensional positional information of an object and an instruction of a desired image processing mode from a user of the services; obtaining the three-dimensional positional information of the object captured in the originally presented image from the object image received; performing image processing on the received image in the instructed image processing mode in accordance with the three-dimensional positional information obtained; and sending out image data obtained as a result of the image processing to the user of the services.

A solution taken by the invention of Claim 20 is an order processing method comprising the steps of: receiving an instruction that specifies a type of a product from a user; sending out an image containing three-dimensional positional information of the product of the specified type to the user; and receiving an order for the product, along with information specifying the size of the product, from the user.

[Embodiments of the Invention]

Hereinafter, embodiments of the present invention will

be described with reference to the drawings.

As used herein, a "range image" is an image representing either distances associated with respective pixels as measured from a camera or depth values in the three-dimensional coordinate system. Each of these distances corresponds to a spherical coordinate r of the spherical coordinate system (r, θ, ϕ), while each of the depth values corresponds to an orthogonal coordinate z of the orthogonal coordinate system (x, y, z). In a first embodiment, the spherical coordinate r will be used. On the other hand, the orthogonal coordinate z will be used in a second embodiment. It should be noted, however, that the spherical coordinate system (r, θ, ϕ) and orthogonal coordinate system (x, y, z) are mutually convertible bidirectionally. Also, by using a pixel location with coordinates (x, y) on a CCD, (r, θ, ϕ) and (x, y, z) can be easily obtained from r and z, respectively, through three-dimensional geometric computations.

(FIRST EMBODIMENT)

Figure **1** is a block diagram of a geometric measuring camera as an image processor according to the first embodiment of the present invention. In Figure **1**, **1** denotes a casing of the camera, **11a** and **11b** denote first and second flash tubes to be light sources, **12** denotes a light source controller for controlling the light emitted from the first

and second flash tubes **11a** and **11b**, **13** denotes an imager, **14** denotes a range image generator for generating a range image containing data about the distances associated with respective pixel locations on the image as measured from the camera from the image captured by the imager **13**, and **15** denotes a color image generator for generating a normal color image from the image captured by the imager **13**. The first and second flash tubes **11a** and **11b** and light source controller **12** together constitutes a light-emitting means **16**. And the imager **13** and light-emitting means **16** together makes up an imaging section **10**. The light-emitting means **16** is formed so as to be removable from the body, i.e., the imaging section **10**.

21 denotes a display panel as a display for presenting an image thereon, **22** denotes a touch panel provided on the surface of the display panel **21**, **23** denotes a shutter release and **24** denotes a recording/reproducing section for recording or reproducing the range image or color image onto/from a storage medium **25**.

A control unit **30** not only controls the operations of the light source controller **12** and imager **13** in response to a signal of the shutter release **23** but also gets the input color image presented on the display panel **21**. Also, the control unit **30** includes an image synthesizer **31** for combining a scale image representing a real size at a specified position on the image with the color image.

In the geometric measuring camera shown in Figure **1**, the imager **13** can capture an image including three-dimensional positional information as in the described prior-art example. This embodiment is characterized in that a range image generated by the range image generator **14** is used as the three-dimensional positional information and a scale image representing a substantially real size can be presented, along with a color image, in accordance with the range image.

Figure **2** is a schematic diagram illustrating the appearance of the geometric measuring camera shown in Figure **1** when viewed from the backside of the camera. The display and touch panels **21** and **22** are stacked on the backside of the camera **1**. By manipulating the touch panel **22**, the user can specify the position on the captured image presented on the display panel **21** at which a real size of the object is presented by a scale image.

It will be described with reference to Figure **3** how the geometric measuring camera of this embodiment works for the presentation of a scale image along with how the user should operate the camera.

First, as shown in Figure **3(a)**, the user takes a picture of an object, e.g., a desk placed in a room, with the geometric measuring camera of this embodiment. In this case, the image obtained includes three-dimensional positional information of respective parts of the desk. Then, the user

has a color image of the desk presented on the display panel **21**.

Next, as shown in Figure **3(b)**, the user specifies his or her desired position on the image at which the real size should be represented by the scale image. For example, the user may select a scale display mode for the display panel **21** (a mode switch, for instance, may be provided for the body of the camera in that case). Then, the user may specify a position **C** (i.e., a point on the upper surface of the desk in the drawing) using the touch panel **22** while watching the screen.

In response, the image synthesizer **31** calculates a real size for the specified position **C** based on the range image thereof, generates a scale image representing the real size obtained and then combines the scale image with the color image. As a result, the scale image **S** representing the real size of the upper surface of the desk is presented at the specified position **C** as shown in Figure **3(c)**. In this case, the scale image **S** is an image representing a ruler with a scale.

The real size may be calculated in the following manner. Based on an Euclidean distance between the specified position **C** with a set of three-dimensional coordinates and the center of the camera lens with a different set of three-dimensional coordinates, the distance **L** between the camera and the object

is derived. The real length $\mathbf{P_x}$ represented by each pixel on the screen for the specified position **C** is given by

$$P_x = L/f \cdot S_x/N_x$$

where f is the focal length of the camera, S_x is a CCD size of the camera and N_x is the number of horizontal active pixels in the CCD. Accordingly, the number $\mathbf{N_p}$ of onscreen pixels equivalent to a real length **X** is given by

$$N_p = X/P_x$$

Thus, the scale image **S** with a length represented by the pixel number $\mathbf{N_p}$ is combined with letters **L** representing the real length **X** (e.g., 20 cm in the drawing) at the specified position **C**.

Also, as shown in Figures **3(d)** and **3(e)**, the user may change the direction and display position of the presented scale image **S** by manipulating, for example, the touch panel **22**. It should be noted that when the scale image **S** presented is directed toward the depth of the screen as shown in Figure **3(d)**, the distance **L** between the camera and the object changes gradually. Accordingly, the scale image **S** will have a non-even scale in that case. Optionally, the processor may also be formed so that the user can select, for example, one of multiple scale images **S** having scales with different widths.

If the user wants to remove the scale image **S** from the screen, then he or she cancels the scale display mode with

the mode switch, for example. As a result, the scale image **S** disappears and only the original color image is left as shown in Figure **3(f)**.

In this manner, according to this embodiment, the user can easily know a real size of an imaged object by reference to the scale image representing the real size on the screen while watching the image. Thus, on looking at an image of an object, e.g., a fish that has been caught or a finding that has been dug up in some remains, the user can know its real size easily without placing any size reference object beside it. Also, if the camera of this embodiment is used as a security camera, then the effect that the user can easily identify a body size of a criminal imaged is obtained.

Further, in this embodiment, the specified position **C** is supposed to be selected using the touch panel **22**. Alternatively, a penlike pointing device that is formed so as to allow the user to specify coordinates on the display panel **21** may be used instead of the touch panel **22**. Cursor keys, press buttons, dial or mouse that are formed so as to allow the user to move a cursor presented on the display panel **21** may also be used. Figure **4** is a diagram illustrating the back of the camera **1** where it is equipped with cursor keys **28**. In Figure **4**, the camera **1** is so formed as to allow the user to move and set the position of a cursor **27** presented on the display panel **21** by using the cursor keys **28**. In this

case, the touch panel **22** is not needed and can be omitted.

Also, as three-dimensional positional information needed in presenting the scale image, data about the distance between the camera and the object, obtained by an auto- or manual focusing controller built in the camera, may also be used. In that case, there is no need to capture an image containing the three-dimensional positional information or to provide the light-emitting means made up of the first and second flash tubes **11a** and **11b** and light source controller **12**. That is to say, any camera can present a scale image so long as the camera includes the auto- or manual focusing controller and a normal imager that can obtain information about the distance between the camera and the object.

Specifically, the real lengths (P_x, P_y) represented by each pixel on the CCD are calculated for a distance **L** by the following equations:

$$P_x = L/f \cdot S_x/N_x$$

$$P_y = L/f \cdot S_y/N_y$$

where L is the distance data obtained by the auto- or manual focusing controller (i.e., the distance between the camera and the object), f is the focal length of the lens, (S_x and S_y) are the sizes of the CCD, and (N_x and N_y) are the numbers of active pixels on the CCD. The subscripts **x** and **y** denote the horizontal and vertical directions, respectively. Then, a scale image is generated based on these lengths (P_x, P_y).

For example, a scale image with a length represented by a pixel number **N** is generated, and the real length $\mathbf{N} \times \mathbf{P_x}$ (in the x direction) or $\mathbf{N} \times \mathbf{P_y}$ (in the y direction) are presented on the screen along with the scale image. Alternatively, a pixel number $\mathbf{R/P_x}$ (in the x direction) or $\mathbf{R/P_y}$ (in the y direction) equivalent to a real length **R** may be obtained to generate a scale image with a length represented by the pixel number and then the length **R** may be presented on the screen along with the scale image.

Also, if the calculation results representing sizes of the object are stored on the storage medium **25** by the recording/reproducing section **24**, the user does not have to remember the calculation results. Furthermore, it is convenient because the information stored on the storage medium **25** can also be utilized by using a personal computer, for example, with functions equivalent to those of the recording/reproducing section **24**. Naturally, the range image, the color image or the synthesized image obtained by combining the scale image with the color image may be stored on the storage medium **25**.

Moreover, where the CCD sizes, focal length of the lens during imaging and distance data of the object obtained by the auto- or manual focusing controller are stored, along with the image data, on the storage medium **25** by the recording/reproducing section **24**, the scale image can also be

presented afterward along with the image already recorded even if no images are being captured.

Also, as shown in Figure **5**, if an image memory **35** is provided to temporarily store the image data received from the imager **13** thereon, pictures can be input to the camera consecutively and then recorded by the recording/reproducing section **24** on the storage medium **25** afterwards. In addition, multiple pictures may be read out from the storage medium **25** onto the image memory **35** and then presented at high speeds.

Further, any image may be used as the scale image so long as it can be a reference representing a real size of an object. Therefore, the shape of the scale image is not limited to a ruler. For example, the scale image may represent baseball bat, golf club or even a row of portraits of characters. Furthermore, the user may also select a type of scale image or change the types of scale images depending on the image to be presented.

Also, according to this embodiment, an object may be presented substantially in its real size on the display panel **21** by utilizing the real size information (i.e., three-dimensional positional information) of the object contained in the image data obtained. This is easily realizable by providing the length $\mathbf{P_x}$ of a single pixel on the display panel **21** to the image synthesizer **31**. Specifically, the number $\mathbf{N_p}$ of pixels equivalent to a real length **X** is given by

$N_p = X/P_x$

Thus, the image synthesizer **31** may scale the image up or down so that when the image is presented on the display panel **21**, its number of pixels corresponding to the real length **X** becomes equal to the number $\mathbf{N_p}$.

It should be noted that the display panel **21** is built in the casing **1** and its size is relatively small. Accordingly, it is impossible to present an image of a big object in its real size. However, by using an external display panel such as a desktop personal computer, even a big object can be presented thereon in its real size. Also, the image representing the object in its real size may be output as a hardcopy using a printer.

(SECOND EMBODIMENT)

A geometric measuring camera as an image processor according to a second embodiment of the present invention has essentially the same configuration as that illustrated in Figures **1** and **2**. But the image synthesizer **31** of this embodiment is different from that of the first embodiment in that it combines a plurality of images together so that the images are in proportion to their actual sizes. That is to say, this embodiment is characterized in that the range images generated by the range image generator **14** are also used as three-dimensional positional information and a synthesized image are obtained by combining multiple images

together so that the images are in proportion to their real sizes based on the three-dimensional positional information and then presented.

It will be described with reference to Figure **6** how the geometric measuring camera of this embodiment works for the combination of multiple images along with how the user should operate the camera. In this embodiment, the case where an image of a desk is combined with an image of the user's room is taken as an example.

First, as shown in Figure **6(a)**, the user takes a picture of an object, i.e., a desk, using the geometric measuring camera of this embodiment. Then, the user gets the image data representing the range and color images of the desk that have been generated from the captured image by the range and color image generators **14** and **15**, respectively, stored on the storage medium **25** by the recording/reproducing section **24**.

At the same time, as shown in Figure **6(b)**, the user also takes a picture of a room in his or her home, for example, using the geometric measuring camera of this embodiment. Then, the user obtains the range and color images of the room and gets the image data stored on the storage medium **25** by the recording/reproducing section **24**.

Next, as shown in Figure **6(c)**, the user reads out the desk image stored and gets the image separated from its background image and then presented on the display panel **21**.

An object image can be separated from the background image by getting presented only a portion of the captured color image that is defined by predetermined three-dimensional coordinates. In the illustrated embodiment, by using the range image, a portion of the color image, which is made up of pixels at respective locations associated with distances falling within a predetermined range from the camera, is cut out as the desk image. For example, if portion of the range image where distance data is more than 2 meters is regarded as belonging to the background, only an image of the object, which is located less than 2 meter away from the camera, i.e., the desk image in this case, can be cut out.

Then, the user specifies an alignment point α on the separated desk image using the touch panel **22.** In the illustrated example, the user wants to place the desk image at a room's wall image portion. Accordingly, the user defines a point α at an upper edge of the desk image as a reference point to be aligned with the room's wall image portion.

Subsequently, as shown in Figure **6(d)**, the user reads out the room image stored and gets the image presented on the display panel **21.** Then, the user specifies another alignment point β using the touch panel **22.** In the illustrated example, the user sets the point β on that image portion representing the room's wall.

Thereafter, as shown in Figure **6(e)**, the user combines

the desk image shown in Figure **6(c)** with the room image shown in Figure **6(d)**. These images are combined together in such a manner that those two specified points α and β are located at the same position within the three-dimensional space. That is to say, a synthesized image representing a room with a desk placed is generated so that the point α is aligned with the point β.

In this case, however, the desk and room images were captured under conditions that the distance between the camera and one of the objects is different from the distance between the camera and the other and these images were captured with the magnification of a zoom lens defined at mutually different values. Accordingly, if these images are combined together as they are, then the resultant synthesized image might not be in proportion to their real sizes. That is to say, there arises a problem that the scale of the desk image might be too large or too small compared to the room image in view of their desired size relationship.

Thus, according to this embodiment, the desk and room images are combined together in such a manner that the images are in proportion to their actual sizes using the range images. Specifically, this process includes the following processing steps.

First, the magnitude of translation (l, m, n) between the specified points α $(x\alpha, y\alpha, z\alpha)$ and β $(x\beta, y\beta, z\beta)$ is

obtained by the following equations:

$$l = x\beta - x\alpha$$

$$m = y\beta - y\alpha$$

$$n = z\beta - z\alpha$$

Next, each of the pixels, making up the desk image shown in Figure **6(c)**, is translated three-dimensionally by the magnitude (l, m, n). Supposing a point $\mathbf{P_1}$ with a set of three-dimensional coordinates should be translated to another point $\mathbf{P_2}$ with a different set of three-dimensional coordinates, the points $\mathbf{P_1}$ and $\mathbf{P_2}$ are represented as:

$P_1 = [x1\ y1\ z1\ 1]$ and

$P_2 = [x2\ y2\ z2\ 1]$, respectively.

And the point $\mathbf{P_2}$ is given by the following equation

$$P_2 = P_1 \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ l & m & n & 1 \end{bmatrix}$$

Next, a two-dimensional desk image to be presented is derived from the three-dimensional coordinates of the translated desk image through perspective transformation. In transforming a point **P** with three-dimensional coordinates (X, Y, Z) onto another point **Q** with coordinates (x, y) on a projection plane as shown in Figure **7**, these coordinates are given by

$x = f \cdot X/Z$ and

$$y = f \cdot Y/Z$$

where f is the focal length of the camera. If the desk and room images were captured at mutually different zooming magnifications, the process of scaling up or down these images (i.e., projection transformation that equalizes the f values in the above equations with each other) may be performed so that these images are presented at the same magnification.

Thereafter, a texture of the desk surface is extracted from the color image and then mapped onto the obtained desk image.

As a result of this process, the desk image is combined with the room image after their scales have been matched to each other as shown in Figure **6(e)**. An image synthesis process like this is performed by the image synthesizer **31**. In other words, according to this embodiment, a separated object image can be combined with a different background image after at least one of the images has its/their scale(s) adjusted to be in proportion to its/their actual scale(s) even if these images were captured at mutually different distances or zooming magnifications. Accordingly, it is possible to obtain a synthesized image that looks as if the object (i.e., the desk in this embodiment) had been imaged at a place (i.e., the room shown in Figure **6(b)** in this embodiment) different from the place where it was actually imaged.

The alignment points α and β are supposed to be freely

specified by the user. Thus, as a result of the image synthesis, a non-realistic image such as an image of a leg of the desk sinking into the floor of the room or an image of the upper surface of the desk forcing itself into a wall of the room, might be generated. Accordingly, in view of these situations, the relative positions of these images to be combined are preferably made definable or modifiable externally. For instance, the camera may be formed so that the user can move the images using the touch panel **22** or a switch provided on the back of the camera.

For example, in a situation where a desk leg has sunk into the floor, the user should manually specify the magnitude of translation toward the top of the screen. In accordance with the user's instructions, the image synthesizer **31** translates the desk image, performs the image synthesis again and then gets the synthesized image presented on the display panel **21** again. The user repeatedly specifies the magnitude of translation while checking back the synthesized image on the screen until the desk image is placed at an almost satisfactory position.

Further, in this embodiment, the perspective transformation is supposed to be carried out for respective points on the desk image. As an approximation, the distance between the point α and the camera and the distance between the point β and the camera may be obtained from the range

images and a two-dimensional scaling process may be performed with a ratio of these distances regarded as the magnification of the desk image, thereby combining the obtained images.

Also, in this embodiment, each of the images is transformed so as to be in proportion to its real size. However, it is easy for the user to intentionally change the scale of the object image. In that case, the image may be scaled up or down two-dimensionally or the object images may be scaled up or down three-dimensionally based on the three-dimensional positional data of the images and then an image to be presented may be obtained through perspective transformation. In other words, the image synthesizer **31** may combine multiple images together in such a manner as to meet a desired size relationship. As used herein, "combining images to meet a desired size relationship" includes combining the images together so that at least one of the images is substantially in proportion to its actual size as described above.

That is to say, the image synthesizer **31** is so formed as to scale up or down at least one of multiple images to be combined. In this case, the scaling process of scaling up or down the image may be performed through the following three-dimensional transformation:

$P_1 = [x1 \ y1 \ z1 \ 1]$

$P_2 = [x2 \ y2 \ z2 \ 1]$

$$P_2 = P_1 \cdot \begin{bmatrix} kx & 0 & 0 & 0 \\ 0 & ky & 0 & 0 \\ 0 & 0 & kz & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In this case, the user can intentionally scale up or down the image of a particular object. Thus, a miniaturized synthesized image can be obtained, for example.

Also, if the camera is tilted during imaging, for example, then the resultant object image (the desk image in this embodiment) might also tilt. In that situation, for example, the desk image should preferably be combined with the room image after the horizontal plane of the desk image has been parallelized to that of the room image by subjecting the desk image to rotational transformation. To realize the rotational transformation like this, the 4×4 matrix in above equation may be replaced with the following matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -m & -n & 1 \end{bmatrix} \begin{bmatrix} \\ R \\ \\ \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & m & n & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} n_x^2 + (1 - n_x^2)\cos\theta & n_x n_y (1 - \cos\theta) + n_z \sin\theta & n_x n_x (1 - \cos\theta) - n_y \sin\theta & 0 \\ n_x n_y (1 - \cos\theta) - n_z \sin\theta & n_y^2 + (1 - n_y^2)\cos\theta & n_y n_z (1 - \cos\theta) + n_x \sin\theta & 0 \\ n_x n_z (1 - \cos\theta) + n_y \sin\theta & n_y n_z (1 - \cos\theta) - n_x \sin\theta & n_z^2 + (1 - n_z^2)\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where R is a rotational matrix. By using this matrix, the image is rotated by θ degrees around the axis passing the

origin and a point (l, m, n). n_x, n_y and n_z are specified as follows:

$$n_x = \cos\alpha$$

$$n_y = \cos\beta \text{ and}$$

$$n_z = \cos\gamma$$

where α, β and γ are angles formed by the axis of rotation with the x-, y- and z-axes, respectively. In a special situation where the image has been rotated around the x-axis,

$$n_x = 1 \text{ and } n_y = n_z = 0$$

Or if the image has been rotated around the y-axis, then

$$n_x = n_z = 0 \text{ and } n_y = 1$$

Or if the image has been rotated around the z-axis, then

$$n_x = n_y = 0 \text{ and } n_z = 1$$

In this manner, the rotational matrix **R** can be modified into a simple form.

Also, as three-dimensional positional information needed in matching the scales of multiple images to each other, data about the distance between the camera and an object, obtained by the auto- or manual focusing controller built in the camera, may also be used. In that case, there is no need to capture an image containing the three-dimensional positional information or to provide the light-emitting means made up of the first and second flash tubes **11a** and **11b** and light source controller **12**. That is to say, any camera can approximately match the scales of the images to each other so long as the

camera includes the auto- or manual focusing controller and a normal imager that can obtain information about the distance between the camera and an object.

Further, in this embodiment, the image is a natural image to be taken with a camera. Alternatively, an image synthesized by a computer graphics process may also be used.

Also, in this embodiment, two types of images, i.e., the desk and room images, are combined. However, even in the case where three or more types of images are combined together, the images can be combined so as to be in proportion to their actual sizes if the above process is performed on each of these images. In those cases, for example, images of multiple objects that were captured in mutually different situations can be combined together on a common background image with their scales matched to each other.

Further, in this embodiment, the alignment points α and β are supposed to be specified using the touch panel **22**. Alternatively, a penlike pointing device formed so as to allow the user to specify coordinates on the display panel **21** may be used instead of the touch panel **22**. And cursor keys, press buttons, dial or mouse, which are formed so as to allow the user to move a cursor presented on the display panel **21** as shown in Figure **4**, may also be used. In that case, the touch panel **22** is not needed and can be omitted.

Furthermore, so long as the CCD sizes, focal length of the lens during imaging and distance data of the objects obtained by the auto- or manual focusing controller are stored, along with the image data, on the storage medium **25** by the recording/reproducing section **24**, images can be combined in an arbitrary manner when the images are presented afterwards even if no images are being captured.

Also, the light-emitting means **16** is so formed as to be removable. Accordingly, the camera may be used with the light-emitting means **16** attached thereto when range images should be captured and the camera may be used with the light-emitting means **16** removed therefrom when normal images should be captured. The light-emitting means **16** may naturally be fixed to the camera.

Furthermore, in the first and second embodiments, a geometric measuring camera, which captures a range image (i.e., depth image) by imaging the light that has been projected from the light-emitting means **16** and then reflected off from the object, is taken as an example to describe the present invention. Alternatively, the present invention is also easily implementable even with a camera utilizing any other depth measuring method. For example, the present invention is applicable to a method of performing triangulation in which images captured by two horizontally arranged cameras are matched stereoscopically or a method of

measuring the time of flight of a laser spot beam that has been projected to scan a space with time, because a depth image can be obtained in accordance with these methods. In short, so long as two-dimensional images and range information can be obtained, the present invention is applicable to any camera no matter how the camera collects the range information.

Moreover, the scope of the application of the present invention is not necessarily limited to cameras. Specifically, the present invention is also implementable as an image processor with no imaging section, e.g., a personal computer, if the apparatus has the function equivalent to that of the image synthesizer **31**. In that case, responsive to a color image and image data, e.g., range image, containing three-dimensional positional information, the image processor either presents a scale image or combine multiple images together with their scales matched through the same process as that exemplified in the first or second embodiment. As a means for inputting user's instructions, mouse, keyboard, track ball, switch and volume and the like are used.

It should be noted that part or all of the functions of the inventive image processor may be implemented using either dedicated hardware or software that includes a program to be executed by a computer. Also, the program, getting the functions of the inventive image processor executed by a

computer either partially or fully, may be recorded on a computer-readable storage medium so that the computer can execute the program recorded on the storage medium.

It should also be noted that, in each of the embodiments, the case where the geometric measuring camera shown in Figure **1** is supposed to take a still picture is described. However, even if the camera can take a moving picture, the present invention is implementable in a similar manner. In that case, a light source that is formed to be capable of projecting two different types of light patterns onto an object alternately and successively as shown in Figure **11** should be used. A method of using a light source like this includes, for example, using a stroboscopic light source that flashes successively, and shaping a laser beam emitted from a semiconductor laser diode into a linear beam through a rod lens and then having its intensity modulated while making the linear beam sweep the object using a galvanic mirror. It should be noted that if moving pictures of objects have been taken, then those images may be naturally combined or presented as they are. It is also possible, for example, to process a still picture contained in the moving picture and then combine the still picture processed with the moving picture to present the combined picture.

(Application to image processing services)

Also, in each of the embodiments, the functions of

imaging, image generation, image synthesis and presentation are all implemented within the casing of the camera **1** and the camera is so formed as to realize all of these functions by itself. However, the same effects are also attainable even if these functions are realized by discrete units with data exchanged between these units. Such an embodiment of the present invention realizes novel image processing services.

Figure **14** is a schematic diagram illustrating the image processing services according to the present invention. In Figure **14**, an imaging section **61** includes the light sources **11a** and **11b**, light source controller **12**, imager **13**, shutter release **23** and control unit **30** shown in Figure **1** and the user of these services takes a picture of an object using the imaging section **61**. The picture taken should be an image containing three-dimensional positional information or at least information from which the three-dimensional positional information can be drawn. Then, the user sends out the taken image, along with his or her desired mode of processing (i.e., a specific processing method such as three-dimensional positional extract operation or CG data computation) or desired specific type of data processed (i.e., a specific content of output data like three-dimensional positional CG data or three-dimensional positional stereoscopic image data), as a request for image processing to a process service server **63** through a network connection section **62a**.

The administrator of the process service server **63** accepts the image processing request via the network connection section **62a** and the network. Then, the administrator gets range and color images generated from the image data, which the user has obtained using the imaging section **61** and then sent out through the network connection section **62a**, by the range and color image generators **14** and **15**, respectively. Next, the administrator makes the image synthesizer **31** perform the image processing on the image in the specified mode or so that the specified type of data can be obtained through the image processing while taking the size of the object image received into account. And the administrator sends out the image data obtained as a result of this image processing to the user of the services via the network. The user checks the received image data using the display section **64**.

In this case, it is believed that the image processing services to be provided may take any of various forms utilizing the three-dimensional positional information, e.g., combination of a scale image with the object image or combination of multiple images as described above. Also, not only the image data transmitted from the user, but also any other image data may be used for combination. For example, the image data of one of well-known objects that have been stored in advance in the storage of the process service

server **63** may be combined with the image data transmitted from the user.

In this case, the user of the services should pay some money to the administrator of the server as the charge for the process services. The charge is may be either transferred to the server administrator's bank account or paid from the user to the server administrator by way of a credit card company.

It should be noted that, in the illustrated example, the image data is supposed to be transmitted via a network. Alternatively, a storage medium on which the image data is stored may be used to accept or deliver the image data either by post or at a shop.

(Application to customization)

Also, by utilizing the technique of combining images together using the geometric measuring camera shown in Figure **1** to meet a desired size relationship as described in the second embodiment of the present invention, acceptance of an order for a product in accordance with the real size information of the product is implementable.

Hereinafter, it will be described with reference to Figure **15** how to fill an order for a customization on a wristwatch as an example. As shown in Figure **15**, a product catalog server **65** including a database on which the three-dimensional structural information, color image information

and price information of respective products are stored is provided on a network.

First, the user takes a picture of his or her wrist using the geometric measuring camera **1**. Next, the user accesses the product catalog server **65** via the network, checks a product catalog using a browser and selects a wristwatch of his or her desired design. In response, an image of the wristwatch of the specified type, containing the three-dimensional positional information thereof, is transmitted to the user from the product catalog server **65**. Then, the image synthesizer **31** of the geometric measuring camera **1** combines the images so that the images are in proportion to their actual sizes as described in the second embodiment. As a result, an image that looks as though the wristwatch is strapped on the user's wrist is presented on the display panel **21**.

Watching the image on the panel **21**, the user determines whether the design or size of the wristwatch fits in with his or her wrist. And if the user likes it, he or she submits an order for the product to an order processing server **66** via the network.

Further, in this case, if the product is customizable, then the user may adjust the size of the wristwatch presented on the panel **21** using switches, for example, thereby setting the wristwatch image to his or her most preferred size.

Then, the image synthesizer **31** identifies the real size of the wristwatch to be ordered based on the magnification at which the wristwatch image is scaled up or down in this case. If the user places his or her order, along with its real size information, with the order processing server **66**, then the customization is easily realized. Furthermore, in that case, not only the size of the wristwatch as a whole but also respective parts, e.g., the dial, strap and winder, of the wristwatch may have their images scaled up or down independently. And the user may design a wristwatch according to his or her preference and may place an order for the wristwatch in accordance with his or her design. Alternatively, if the three-dimensional positional structural data of the parts is prepared on the catalog server **65**, then the user can combine the images of multiple parts together, present an image of a wristwatch of his or her desired design on the display panel **21** and then place an order for the wristwatch with its parts that make up the wristwatch, color, size, etc. specified. The charge for this order may be paid by any of currently available paying methods. For example, the charge may be paid by credit card, transferred to a specified bank account or paid when the user receives the product.

It should be noted that the above processing method is applicable to customization on various products, not just

wristwatches. For example, in the case where a piece of furniture, e.g., a desk, is customized, first, the user takes a picture of a room in his or her home using the geometric measuring camera **1**. Next, the user accesses the catalog server **65** via the network, views a product catalog using a browser and selects his or her desired furniture. Then, the image synthesizer **31** combines the images so that they are in proportion to their actual sizes as shown in Figure **6(e)** and present the combined image on the display panel **21**. Watching the image on the panel **21**, the user may freely change the scale or position of the furniture using switches. And when the user places an order, he or she may send the ID number of the furniture selected, along with the magnification at which the images are scaled up or down in this case or the real size information of the furniture, to the order processing server **66**.

It should be noted that, in the above description, the size of a product is supposed to be freely set by the user. Alternatively, several different sizes may be prepared for a single product on the receiving end so that the user may select his or her desired size from these. In that case, the product is not totally customizable but can be of just a limited number of sizes. Thus, the costs of developing products can be cut down on the receiving end.

[Effects of the Invention]

As described above, a scale image representing a real size of an object is combined with an original image to present a synthesized image according to the present invention. Thus, on looking at the image presented, the user can know instantly the real size of the object imaged. In addition, multiple images can be combined together so as to meet a desired size relationship, i.e., so that the images are in proportion to their actual sizes, to present the synthesized image. Accordingly, the user can quickly watch a virtual image of an object against a different background, for example.

[Brief Description of the Drawings]

[Figure **1**]

A block diagram of a geometric measuring camera as an image processor according to first and second embodiments of the present invention.

[Figure **2**]

A diagram illustrating the back appearance of the geometric measuring camera shown in Figure **1**.

[Figure **3**]

(a) through **(f)** are diagrams illustrating how the camera of the first embodiment operates.

[Figure **4**]

A diagram illustrating the back of a camera where it is equipped with cursor keys.

[Figure **5**]

A block diagram of a camera where it is equipped with an image memory.

[Figure **6**]

(a) through **(e)** are diagrams illustrating how a camera according to the second embodiment operates.

[Figure **7**]

A diagram illustrating perspective transformation.

[Figure **8**]

A diagram illustrating a basic arrangement for a rangefinder that can capture a range image.

[Figure **9**]

(a) is a diagram illustrating an arrangement for the light sources shown in Figure **8** and **(b)** is a plan view illustrating the arrangement shown in **(a)**.

[Figure **10**]

A diagram illustrating light patterns formed by the light sources shown in Figure **9**.

[Figure **11**]

A graph illustrating a relationship between the light intensity and the angle of projected light for the light patterns shown in Figure **10**.

[Figure **12**]

A graph illustrating a relationship between the intensity ratio of the light and the angle of projected light

for the range α shown in Figure **11**.

[Figure **13**]

A conceptual diagram illustrating how the three-dimensional coordinates are obtained by the rangefinder.

[Figure **14**]

A conceptual diagram of image processing services according to the present invention.

[Figure **15**]

A conceptual diagram of a system for implementing an order processing method according to the present invention.

[Description of the Reference Characters]

10 imaging section

14 range image generator

16 light-emitting means

21 display panel (display)

22 touch panel (input means)

28 cursor keys (input means)

31 image synthesizer

S scale image

[Name of the Document] Abstract

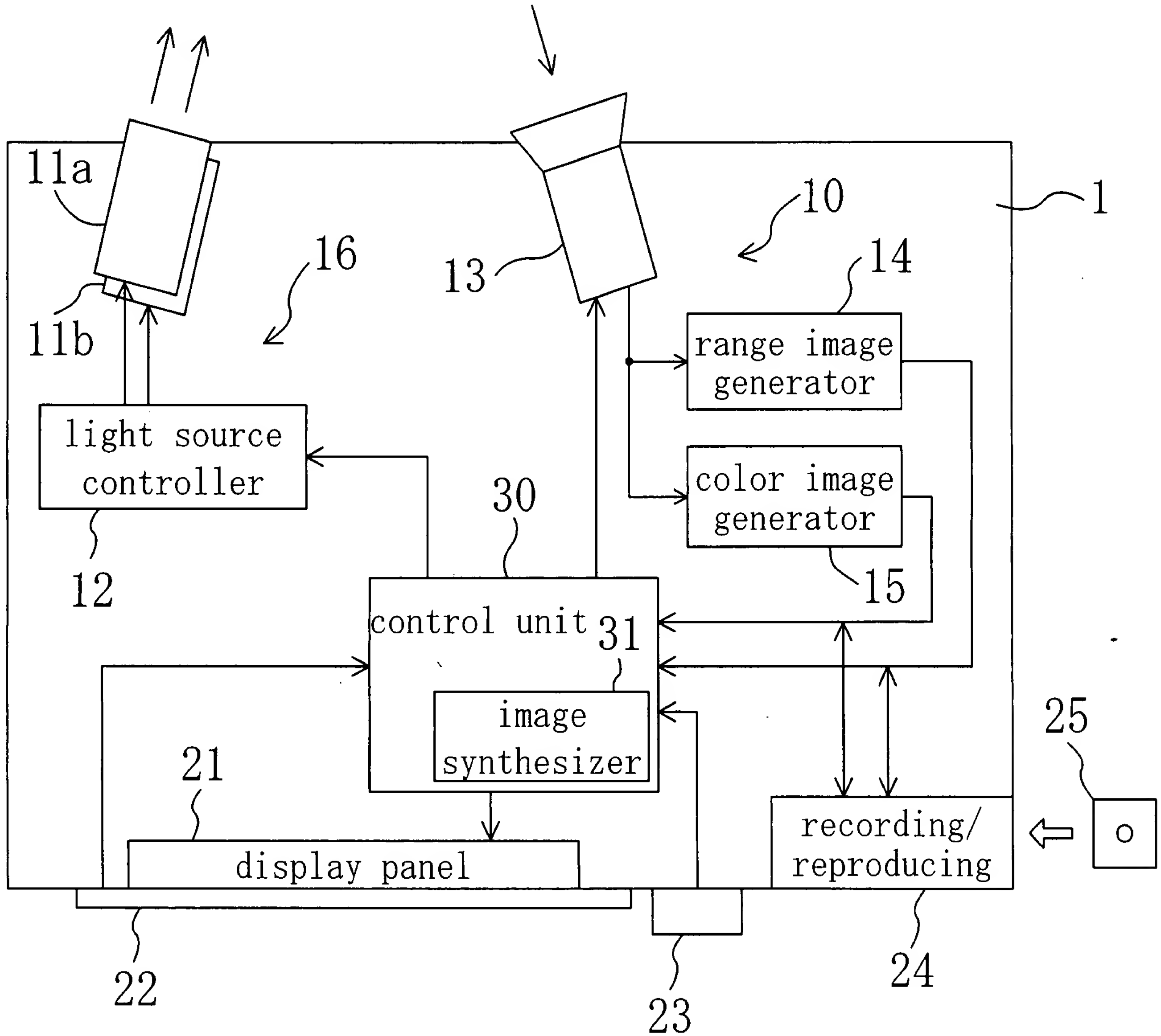
[Abstract]

[Problem] To realize highly convenient and interesting functions for the user in processing an image containing three-dimensional positional information of an object.

[Solution] An image of an object (desk) is presented on a display panel **21**. This image contains three-dimensional positional information of the object. When the user specifies a point **C** on the screen **(b)**, a real size of the object at the specified point **C** is obtained in accordance with the three-dimensional positional information. Thereafter, a scale image **S** substantially representing the real size obtained is generated and then combined with the object image into a synthesized image for presentation **(c)**. The user is allowed to freely change the direction or location of the scale image **S** **(d)** and **(e)**.

[Selected Figure] Figure **3**

FIG. 1



11a
11b
16
13
10
1
14
range image generator
color image generator
15
light source controller
12
30
control unit
31
image synthesizer
21
display panel
22
23
24
recording/ reproducing
25

FIG. 2

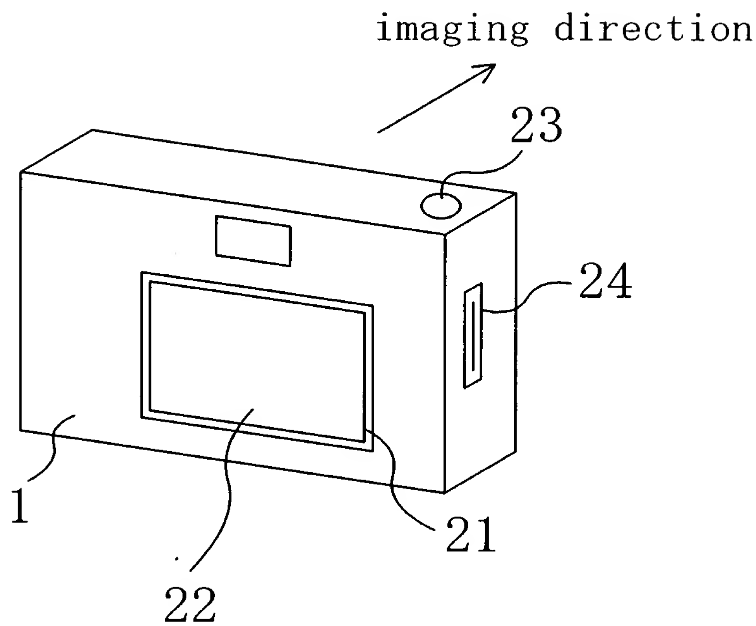


FIG. 3A

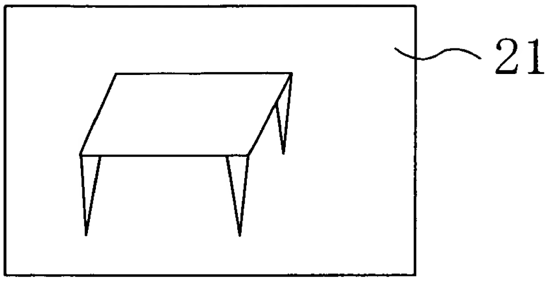


FIG. 3D

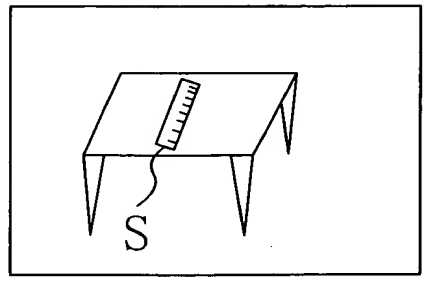


FIG. 3B

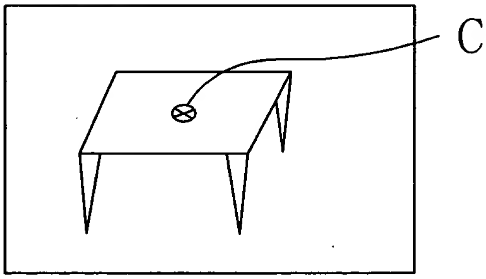


FIG. 3E

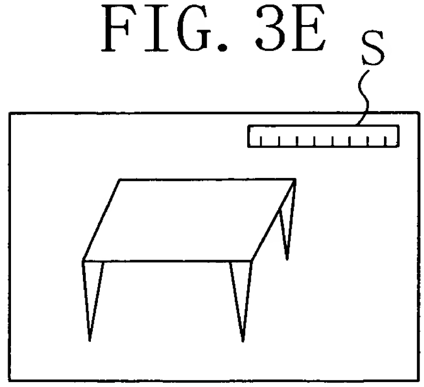


FIG. 3C

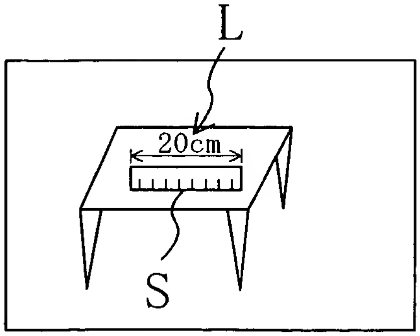


FIG. 3F

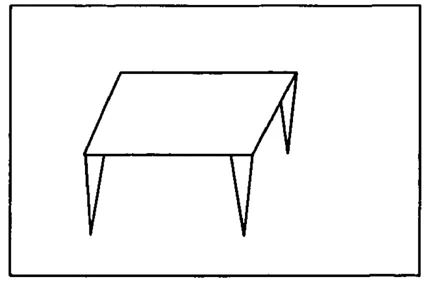


FIG. 4

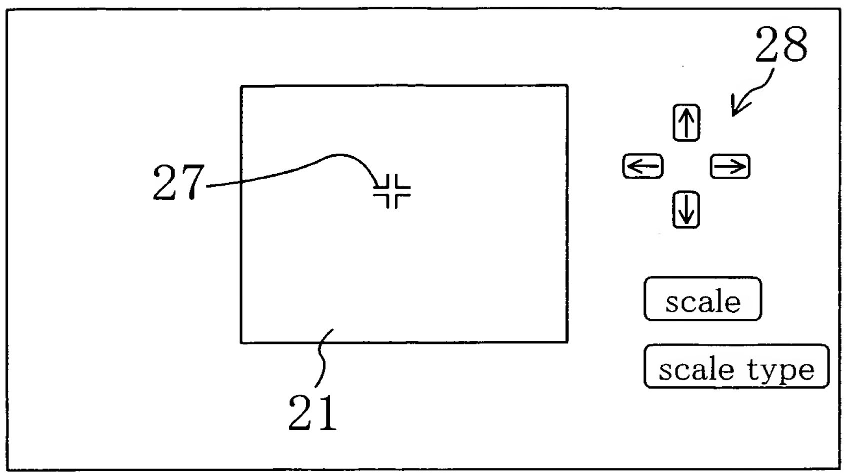


FIG. 5

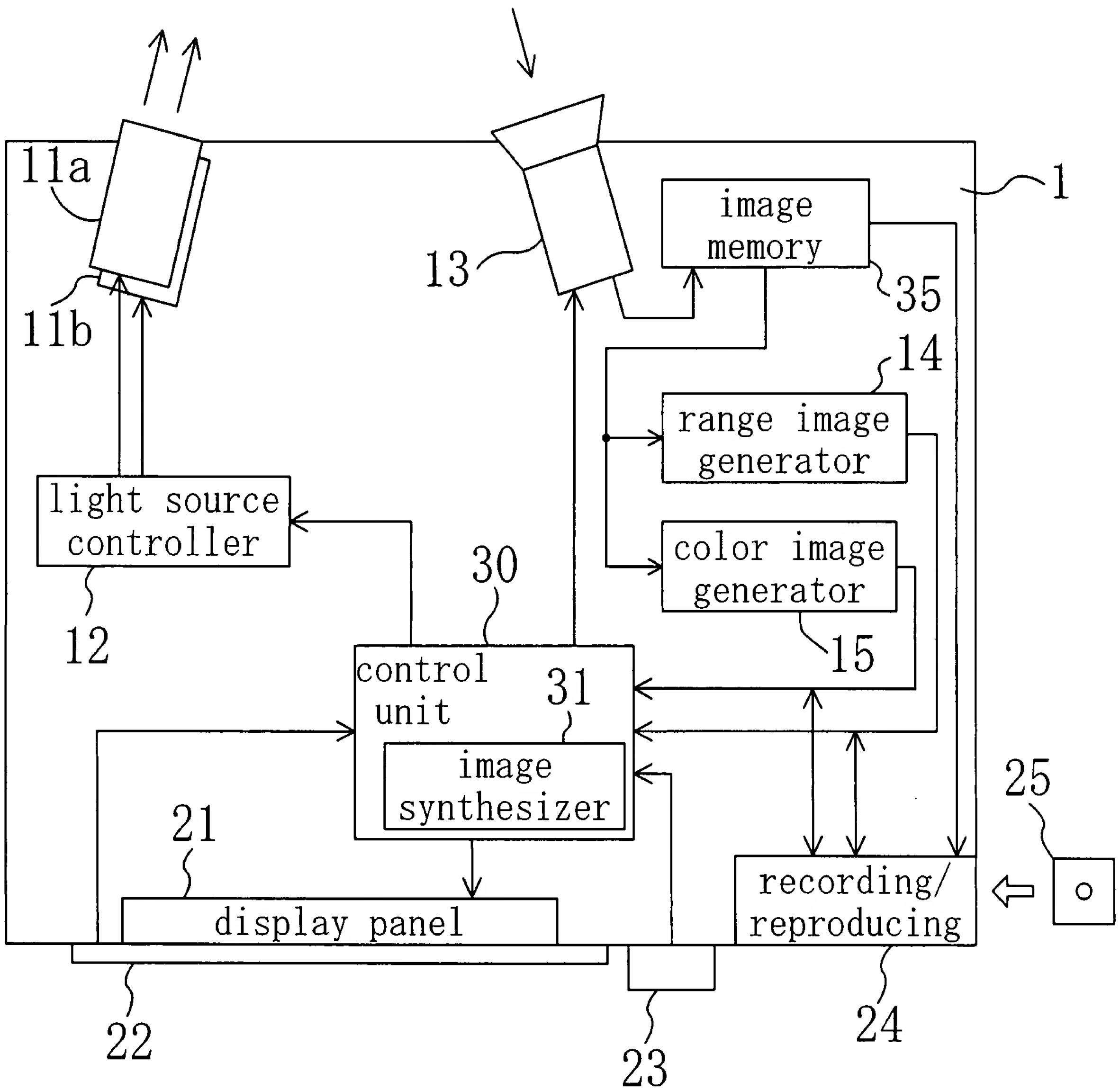

11a
11b
light source controller
12
13
image memory
35
1
14
range image generator
color image generator
15
30
control unit
31
image synthesizer
21
display panel
22
23
recording/ reproducing
24
25

FIG. 6A

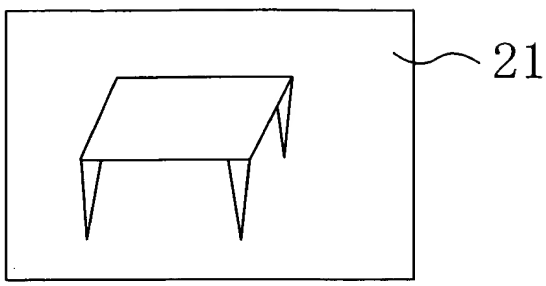


FIG. 6D

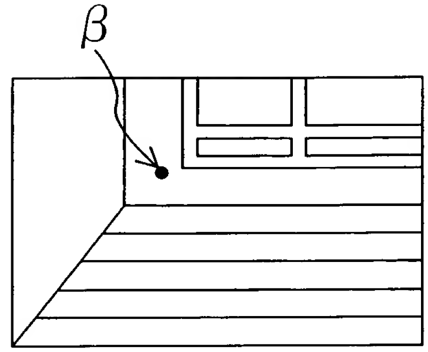


FIG. 6B

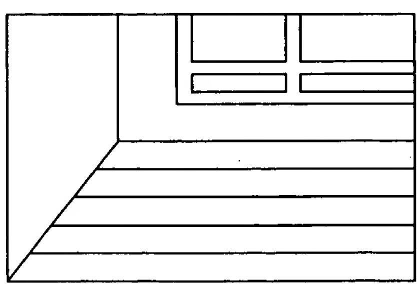


FIG. 6E

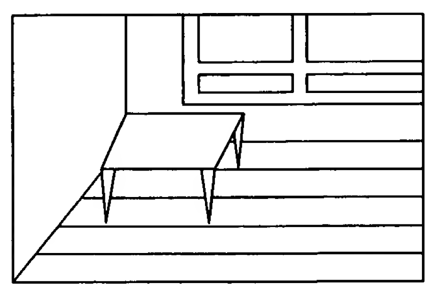


FIG. 6C

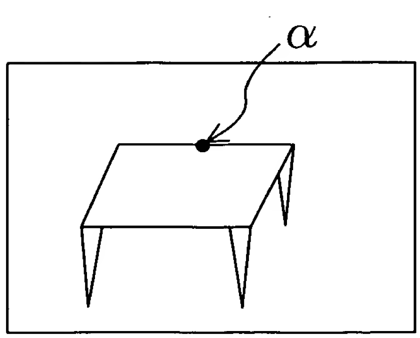


FIG. 7

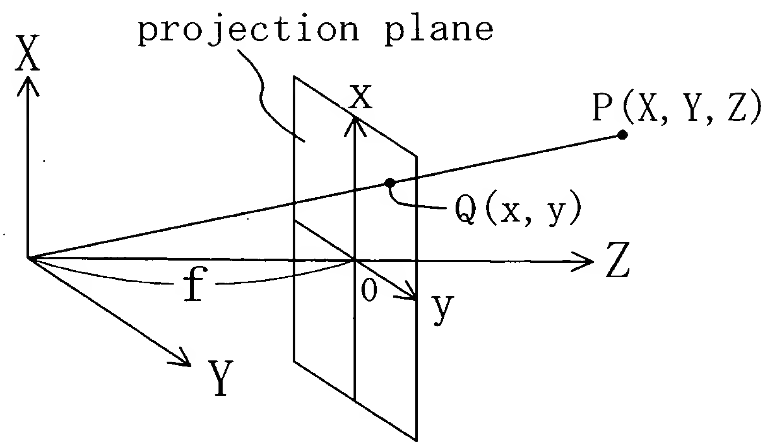


FIG. 8

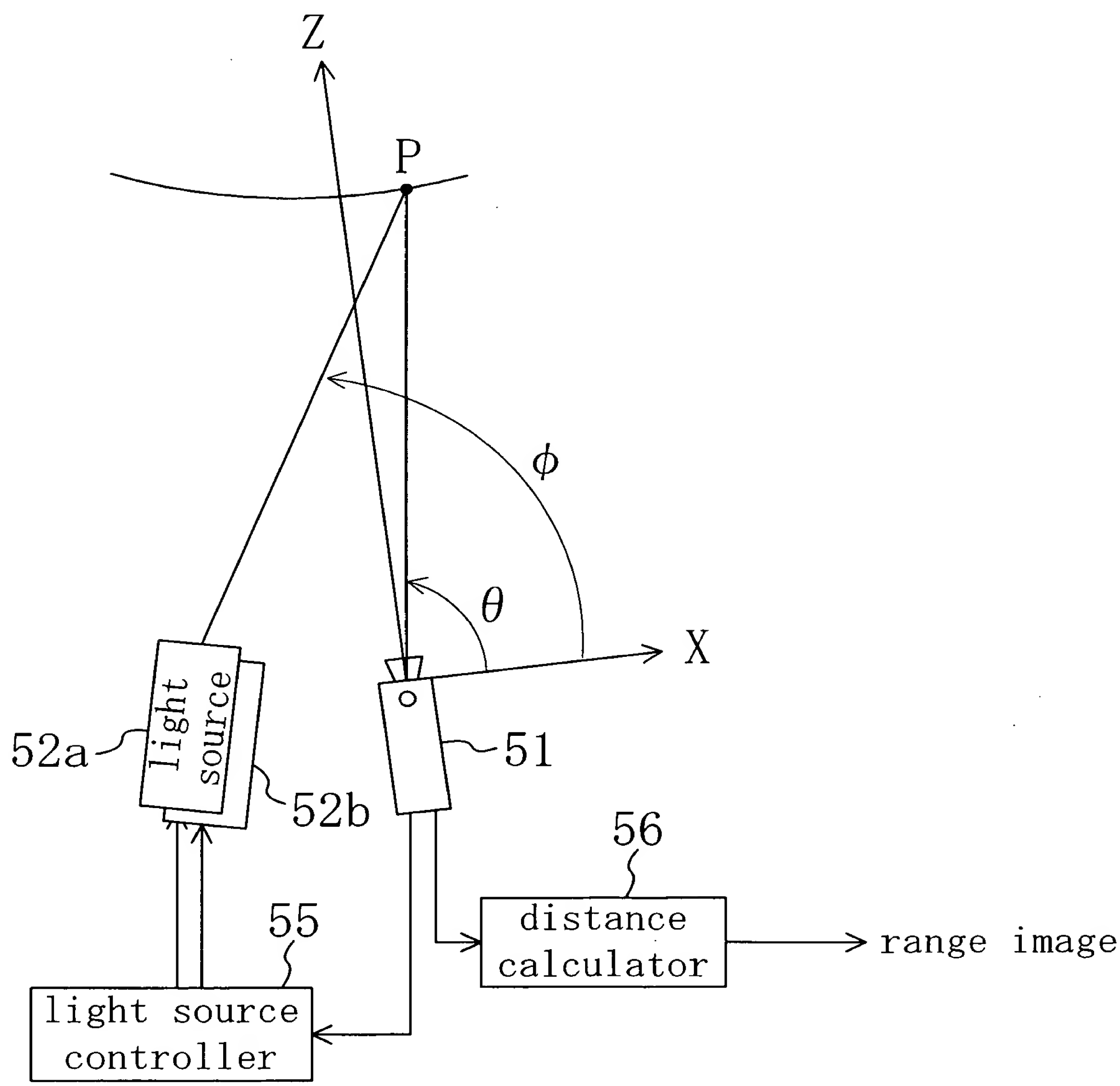

Z
P
φ
θ
X
light source
52a
52b
51
56
distance calculator
range image
55
light source controller

FIG. 9A

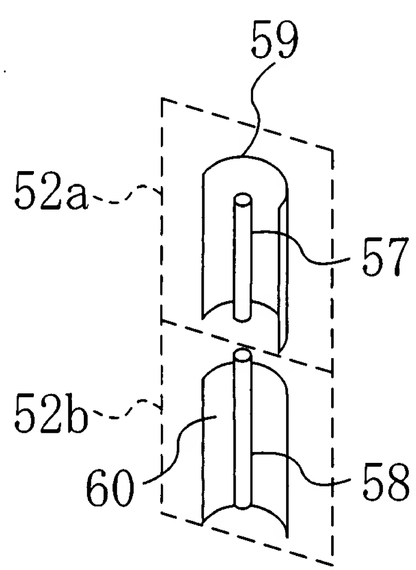


FIG. 9B

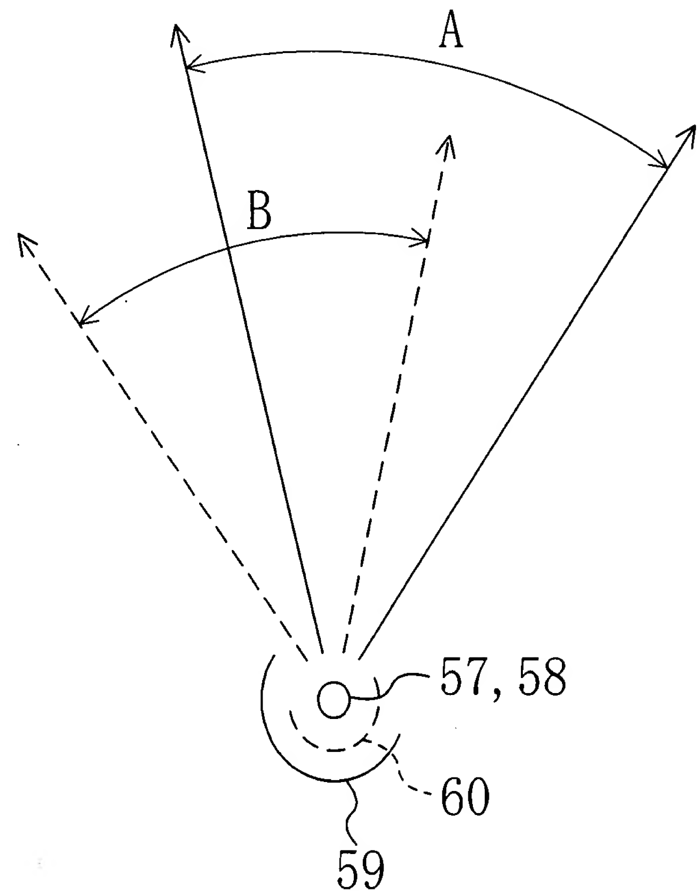


FIG. 10

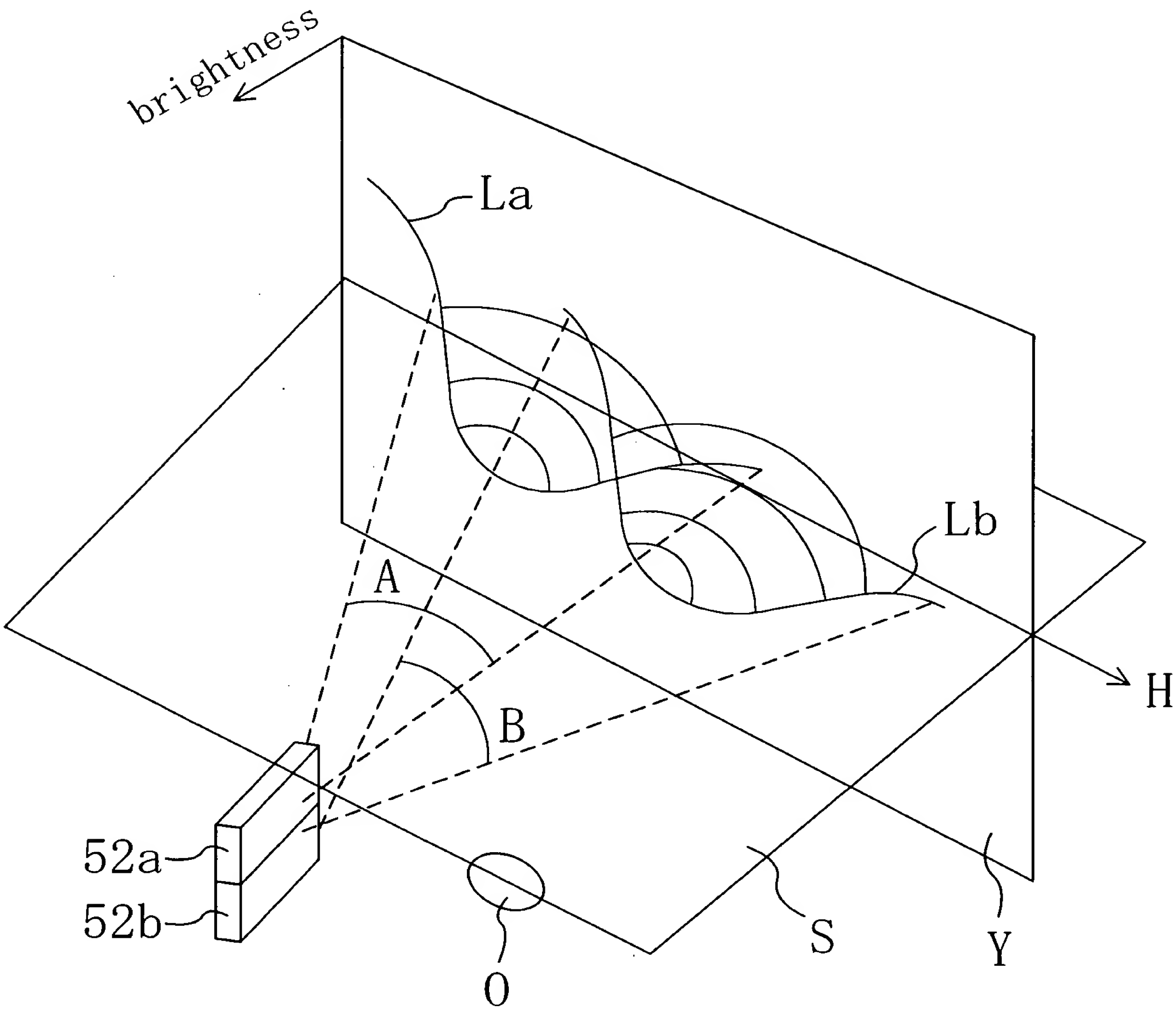
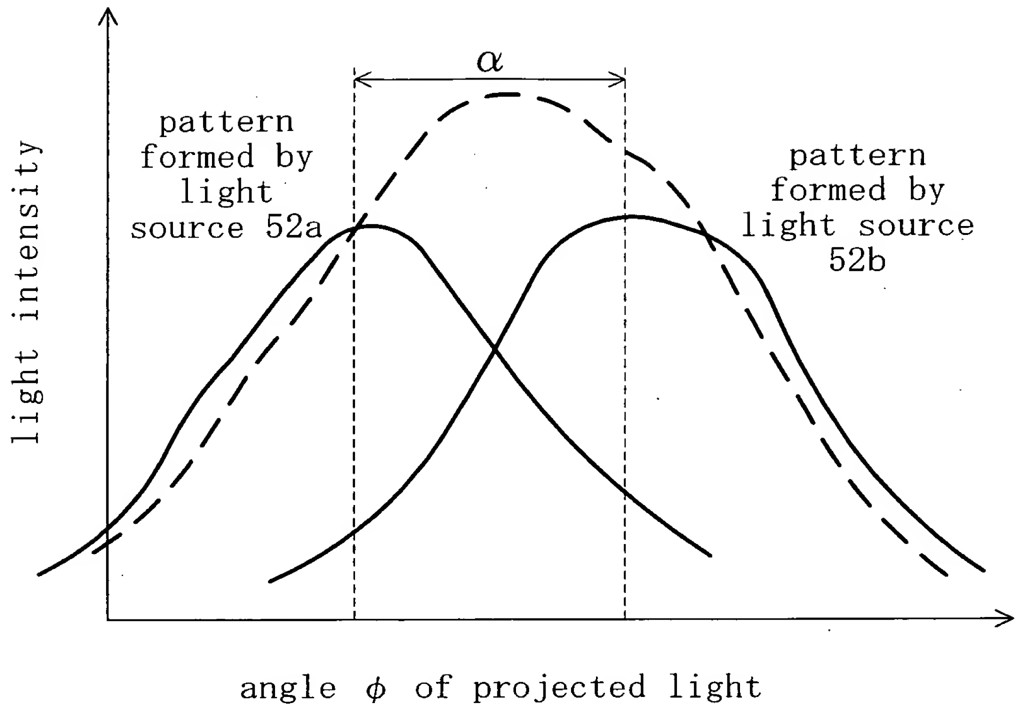
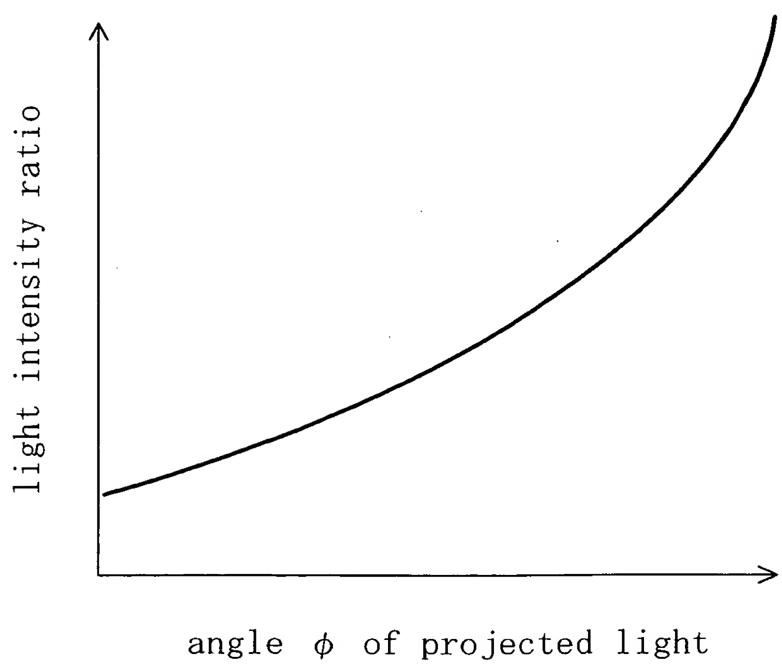


FIG. 11



α
pattern formed by light source 52a
pattern formed by light source 52b
light intensity
angle φ of projected light

FIG. 12



light intensity ratio
angle φ of projected light

FIG. 13